Cell, Vol. 108, 453–463, February 22, 2002, Copyright ©2002 by Cell Press

The RNA Polymerase II Machinery: Structure Illuminates Function

Review

Nancy A. Woychik[1] and Michael Hampsey[2,3]
[1]Department of Molecular Genetics
and Microbiology
[2]Department of Biochemistry
Robert Wood Johnson Medical School
Piscataway, New Jersey 08854

Essential components of the eukaryotic transcription apparatus include RNA polymerase II, a common set of initiation factors, and a Mediator complex that transmits regulatory information to the enzyme. Insights into mechanisms of transcription have been gained by three-dimensional structures for many of these factors and their complexes, especially for yeast RNA polymerase II at atomic resolution.

Overview

RNA polymerase II (RNAP II) catalyzes DNA-dependent synthesis of mRNA but is unable to initiate promoter-dependent transcription or respond to transcriptional regulatory proteins in the absence of other factors. A breakthrough in our understanding the mechanism of transcription initiation followed the discovery that purified mammalian RNAP II would selectively initiate transcription from template DNA when supplemented with a crude cell extract (Weil et al., 1979). This led to the fractionation and subsequent identification of the general transcription factors (GTFs), defined by their requirement for accurate initiation by RNAP II in vitro (Orphanides et al., 1996).

Although the GTFs are sufficient for accurate initiation by RNAP II, transcriptional stimulation in response to promoter-specific activators requires additional factors. Many different factors, from TFIID to chromatin modifiers, have been identified as transcriptional "coactivators," important for activation either in vitro or in vivo. However, most of these are not general factors required for the expression of all RNAP II genes. The most universal cofactor that serves to transduce regulatory information between gene-specific transcription factors and the core RNAP II machinery is a large, modular complex known as Mediator (Myers and Kornberg, 2000). Consistent with its role as a conduit of regulatory information, Mediator can be isolated in a complex with RNAP II, and specific subunits of Mediator interact directly with diverse transcriptional regulatory proteins.

An extraordinary aspect of the RNAP II transcriptional machinery is the extent to which its constituent factors are conserved among eukaryotic organisms. RNAP II is composed of twelve subunits whose sequences are conserved among phylogenetically diverse organisms (Woychik, 1998). The same complement of GTFs has also been identified among disparate organisms, with one-to-one correspondence between each of the subunits. Despite considerable variation in subunit composition, structurally related Mediator complexes have been identified as essential coactivators in mammalian and yeast systems (Hampsey and Reinberg, 1999). Accordingly, the RNAP II transcriptional machinery is conserved through evolution, allowing for the experimental advantages of different organisms to be exploited with the common goal of understanding fundamental aspects of transcription.

The subject of this review is the general RNAP II transcriptional machinery, defined as RNAP II, the GTFs, and Mediator. We summarize recent developments and emphasize the structural advances stemming from site-specific protein-DNA crosslinking studies, electron microscopy, two-dimensional electron crystallography, NMR, and X-ray crystallography. Novel structures, ranging from fragments of individual subunits to the remarkable RNAP II complex at atomic resolution, define the molecular architecture of this machinery.

General Transcription Factors

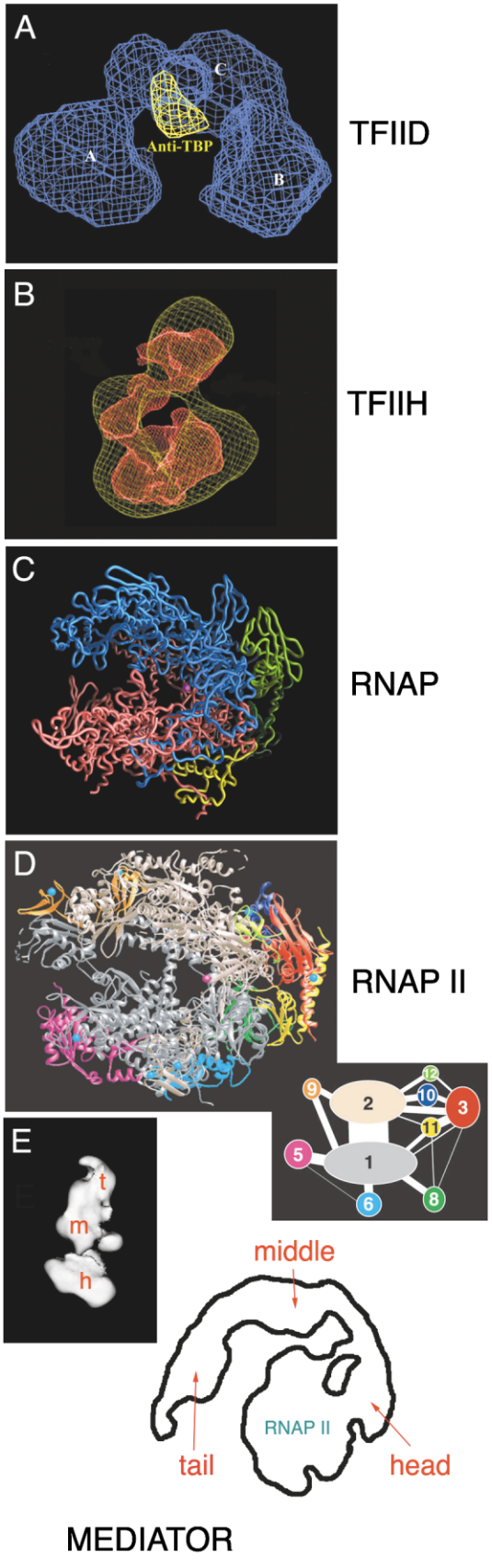
Transcription of protein-encoding genes requires assembly of a preinitiation complex (PIC), composed of template DNA, RNAP II, and five GTFs. PIC assembly is nucleated by binding of the TBP subunit of TFIID to the TATA box, followed by the concerted recruitment of TFIIB, a complex of unphosphorylated RNAP II with TFIIF, TFIIE, and TFIIH. After promoter melting and transcription initiation, the CTD domain of the largest RNAP II subunit is phosphorylated, an event that facilitates promoter clearance and progression into the elongation phase of transcription. Following termination, a phosphatase recycles RNAP II to its unphosphorylated form, allowing the GTFs and RNAP II to initiate another round of transcription (Reinberg et al., 1998).

TFIID

TFIID is composed of TBP plus about ten TBP-associated factors (TAFs). TBP is a single polypeptide that sits astride the TATA box as a molecular "saddle," inducing a sharp bend in the DNA (Nikolov and Burley, 1997). TBP had been regarded as a universal transcription factor, essential for initiation by RNAP I, II, and III. This doctrine has now been challenged by the recent discovery of a metazoan family of TBP-like transcription factors (TLF/TRF2) that bind to DNA sequences different from classical TATA boxes (Dantonel et al., 1999). The physiological role of the TLFs is not yet well defined, although TBP and TLFs have been shown to exert differential effects on developmental gene expression (Holmes and Tjian, 2000; Zhang et al., 2001).

The TAF subunits of TFIID were initially defined as coactivators, mediating the interaction between transcriptional activators and the core machinery (Verrijzer and Tjian, 1996). The TAF dependence of transcriptional activation, however, did not appear to be universal. Activation in a yeast in vitro transcription system was dependent upon a coactivator complex (Mediator) that is devoid of TAFs (Myers and Kornberg, 2000). Subsequent experiments demonstrated that while yeast TAFs are essential for cell growth, several TAFs are dispensable

[3]Correspondence: michael.hampsey@umdnj.edu



for activation in vivo. Among the genes whose expression was affected, TAF dependence appeared to be specified by core promoter elements rather than by upstream regulatory sequences (Green, 2000). Metazoan TAFs have also been shown to bind core promoter elements, including the initiator region (Inr) and the downstream promoter element (DPE). These results do not exclude a role for TAFs as coactivators, though, especially in higher eukaryotes. Indeed, certain TAFs are critical tissue-specific coactivators, presumably involved in the recruitment and stabilization of TFIID at core promoters by gene-specific activators (Freiman et al., 2001, and references therein).

The primary structures of several TAFs exhibit sequence similarity to histones. This sequence similarity extends to structural similarity: a cocrystal structure of *Drosophila* TAF42 and TAF62 revealed a configuration comparable to the histone H3-H4 tetramer (Burley and Roeder, 1996). Although no TAF homolog of histone H2A has been identified, a human H2B-like TAF was found, leading to the proposal that TFIID forms a histone-like octamer complex. Yeast Taf61, Taf17, and Taf60 exhibit structural similarity to histones H2B, H3, and H4, respectively, and several other yeast TAFs also include histone-fold motifs. Indeed, a reconstituted TAF octamer was recently reported, tentatively consisting of a central Taf17-Taf60 tetramer flanked on either side by Taf61-Taf48 dimers, a quaternary arrangement comparable to the histone octamer (Selleck et al., 2001). Whether this TAF octamer exists within native TFIID or binds DNA to form a structure similar to the histone octamer in the nucleosome has not been reported.

In contrast to the gene-specific requirement for several of the yeast TAFs, a more general requirement has

Figure 1. Three-Dimensional Structure of Selected Components of the RNAP II Transcription Machinery

(A) TFIID at 35 Å resolution. The blue mesh represents the three-lobed (A, B, C) TFIID structure. The yellow mesh approximates the position of TBP based on the differential density between TFIID and TFIID bound to a TBP antibody. Reprinted with permission from Andel et al. (1999). Copyright 1999 American Association for the Advancement of Science.

(B) TFIIH at 38 Å resolution. The red mesh represents the 18 Å yeast core TFIIH structure, superimposed onto the 38 Å human holo-TFIIH structure.

(C) *T. aquaticus* RNAP at 3.3 Å. The α carbon backbone is depicted with the following color scheme: α^{I}, light green; α^{II}, dark green; β, cyan; β', pink; ω, yellow; Zn^{2+}, light green sphere; and Mg^{2+}, magenta sphere.

(D) Yeast RNAP II at 2.8 Å. Individual subunits are colored; color and interaction key is provided. The thickness of the white lines approximates the relative amount of surface area buried in the interface between subunits. Reprinted with permission from Cramer et al. (2001). Copyright 2001 American Association for the Advancement of Science.

(E) Yeast Mediator. On the left, 30–35 Å structure of yeast Mediator; the regions comprising head (h), middle (m), and tail (t) modules are depicted. Reprinted with permission from Dotson et al. (2000). Copyright 2000 National Academy of Sciences, U.S.A. Upon interaction with RNAP II, Mediator adopts an extended conformation where the head, middle, and tail domains become more apparent. Bottom right, outline of the yeast Mediator-RNAP II complex depicting the proposed locations of the head, middle, and tail domains relative to RNAP II. Adapted from Dotson et al., 2000.

been reported for the histone-like TAFs (Hahn, 1998). These TAFs are not specific to TFIID, but are also components of the SAGA histone acetyltransferase complex. SAGA includes Taf17, Taf60, and Taf61 but not the putative H2A counterpart, Taf48. The SAGA counterpart of Taf48 could be Ada1, though; it contains a histone-fold motif and forms a heterodimer with Taf61, suggesting the formation of a SAGA-specific histone-like octamer structure (Selleck et al., 2001). TAF-containing complexes other than TFIID, including TFTC, STAGA, and PCAF, have also been identified in higher organisms (Brand et al., 1999, and references therein). Accordingly, the general requirement for histone-like TAFs is likely to reflect their role as core components of multiple complexes.

The molecular architecture of human TFIID and TFTC complexes has been determined by electron microscopy at 35 Å resolution (Figure 1A; Andel et al., 1999; Brand et al., 1999). TFIID is a lobular, horseshoe-shaped structure organized around a solvent-accessible groove that could accommodate a double-stranded DNA molecule. Two different complexes, corresponding to open and closed conformations, were identified, suggesting that TFIID acts like a molecular clamp to bind DNA. Interestingly, none of the lobular domains appears large enough to accommodate all of the histone-like TAFs, arguing against the presence of a histone-like octamer within TFIID. Instead, histone-fold pairs might be present in each subdomain, forming interfaces for protein-protein interactions. The architecture of the TBP-free TFTC complex is similar to TFIID, albeit larger and composed of five lobes.

TFIIB

TFIIB enters the PIC subsequent to formation of the TBP-DNA complex and as a prerequisite for RNAP II binding. TFIIB is a single polypeptide that includes an N-terminal zinc binding domain (nTFIIB), a core domain that encompasses the C-terminal two-thirds of the molecule (cTFIIB), and a phylogenetically conserved sequence that links the two domains. Human cTFIIB includes a helix-turn-helix (HTH) motif that binds the BRE, a sequence present in a subset of promoters immediately upstream of the TATA box (Lagrange et al., 1998). The BRE was recently reported to repress basal transcription, with activator-mediated disruption of the BRE-TFIIB interaction as a proposed mechanism of gene activation (Evans et al., 2001). Neither a comparable HTH motif nor evidence of sequence-specific DNA binding has been reported for yeast TFIIB.

X-ray crystal structures for cTFIIB in DNA-TBP-TFIIB ternary complexes have defined specific TBP-TFIIB and TFIIB-promoter DNA contacts (Bell et al., 1999; Tsai and Sigler, 2000, and references therein). cTFIIB binds the promoter through base-specific contacts in the major groove upstream (BRE) and in the minor groove downstream of the TATA box. This asymmetric binding by TFIIB is likely to account for the unidirectional assembly of the PIC and direction of transcription. An NMR solution structure of cTFIIB revealed notable differences with the crystal structures, suggesting that TFIIB undergoes a conformational change upon entry into the PIC. Furthermore, the cTFIIB conformation is altered by interaction with either nTFIIB or the activation domain of VP16 (Hayashi et al., 1998). This apparent plasticity of TFIIB might underlie earlier reports of activator-induced isomerization of TFIIB and is consistent with a recent solution structure of full-length TFIIB (Hawkes et al., 2000; Grossman et al., 2001). An NMR structure for the N-terminal zinc binding domain has been solved for archaeal TFIIB, revealing a zinc ribbon, comparable to the structure found in the elongation factor TFIIS (Zhu et al., 1996).

The yeast gene encoding TFIIB was initially identified based on mutations that alter start site selection; similar defects in human TFIIB also shift initiation (Hawkes and Roberts, 1999, and references therein). These effects are promoter specific, conferred by sequences in the vicinity of the start sites (Faitar et al., 2001). Although critical for PIC assembly, TFIIB also plays a postrecruitment role in transcription. Moreover, TFIIB defects that affect transcription subsequent to PIC assembly are the same as those that alter the accuracy of initiation (Cho and Buratowski, 1999; Ranish et al., 1999).

Electron crystallography of a TFIIB-RNAP II complex, in combination with the structure of the DNA-TBP-TFIIB ternary complex, suggests a mechanism for start site selection: TFIIB appears to bridge the interaction between TBP and RNAP II such that the DNA template need only follow a straight path from the TATA box to position the start site in the active center of RNAP II (Leuther et al., 1996). This is an attractive model that would account for the conserved spacing between TATA and the start site at many promoters. It is not clear from this model, though, how mutations in either TFIIB or RNAP II affect start site selection, or how yeast RNAP II is able to initiate transcription from multiple start sites within specific promoters. Despite the critical importance of the TFIIB-RNAP II interface, neither the domains nor specific residues that define this interaction have been identified.

TFIIF

TFIIF was initially identified based on its physical association with RNAP II and its requirement for accurate initiation. Binding of RNAP II-TFIIF stabilizes the DNA-TBP-TFIIB ternary complex and is a prerequisite for entry of TFIIE and TFIIH into the PIC. TFIIF is a heterotetramer composed of two large (TFIIFα/RAP74) and two small (TFIIFβ/RAP30) subunits. Both subunits are multiple-domain polypeptides that play distinct roles in initiation, elongation, and regulation of the Fcp1 CTD phosphatase activity (see below).

Site-specific protein-DNA crosslinking experiments have defined contact points of TFIIF with template DNA. RAP30 interacts with template DNA on either side of the TATA box, whereas RAP74 interacts only downstream of TATA (Kim et al., 1997). An independent study reported similar contacts, except for the claim that RAP74 also binds promoter DNA upstream of TATA (Forget et al., 1997). Electron micrograph images of TBP-TFIIB-TFIIF-RNAP II-template complex are consistent between the two studies and suggest that promoter DNA wraps around RNAP II in the PIC. The controversial RAP74 upstream contact points are integral to a model proposed for TFIIF-mediated isomerization of the PIC

(Robert et al., 1998). Higher-resolution images of the TBP-TFIIB-TFIIF-RNAP II-template complex that would depict the topology of TFIIF are not yet available.

An X-ray crystal structure of a heterodimer composed of N-terminal fragments of TFIIF defined a novel "triple barrel" dimerization fold (Gaiser et al., 2000). NMR and X-ray structures for the C-terminal domains of RAP30 (Groft et al., 1998) and RAP74 (Kamada et al., 2001), respectively, revealed remarkably similar tertiary structures, despite limited sequence similarity between the two subunits. These structures form winged helices, similar to the winged helix of the H5 linker histone. Both linker histone and RAP30 bind DNA without sequence specificity but with preference for nonlinear DNA conformations. This could explain the TBP dependence of RAP30 binding to promoter DNA. Furthermore, the two RAP30 winged helices within the TFIIF heterotetramer could account for the crosslinking pattern of RAP30 on either side of the TATA box. As such, RAP30 might be a "condensation factor" (Groft et al., 1998) that could explain the compaction of promoter DNA around RNAP II observed in the electron micrograph images of the PIC (Kim et al., 1997; Forget et al., 1997). The electrostatic properties of the winged helix of RAP74 are substantially different from those of RAP30. Rather than being a DNA binding domain, the RAP74 winged helix has been proposed to interact with the Fcp1 CTD phosphatase (Kamada et al., 2001).

TFIIE

Like TFIIF, TFIIE is a heterotetramer containing two large (TFIIEα) and two small (TFIIEβ) subunits. TFIIE affects late events in PIC assembly, including recruitment of TFIIH and subsequent regulation of TFIIH activities. TFIIE and TFIIH are required for ATP-dependent formation of the open promoter complex prior to formation of the first phosphodiester bond. TFIIE, TFIIF, and TFIIH also cooperate to suppress promoter-proximal stalling, thereby facilitating early events in the transition of RNAP II to productive elongation (Dvir et al., 2001).

The core domain of human TFIIEβ also forms a winged helix motif, the third example of a winged helix in the core RNAP II transcriptional machinery (Okuda et al., 2000). In contrast to the winged helix of RAP74, the winged helices of TFIIEβ and RAP30 are proposed to participate in nonspecific DNA binding. TFIIE forms promoter contacts in and immediately downstream of the transcription bubble region, without altering protein-DNA interactions by RNAP II or the other GTFs (Kim et al., 2000). These TFIIE-promoter contacts are consistent with electron crystallography images of TFIIE with RNAP II, placing TFIIE near the active center of the enzyme (Leuther et al., 1996).

TFIIH

TFIIH is the largest and most complex of the GTFs, consisting of nine subunits with a molecular mass comparable to that of RNAP II. TFIIH is the only GTF with defined enzymatic activities, including two ATP-dependent DNA helicases of opposite polarity (XPB and XPD) and a cyclin-dependent protein kinase (cdk7-cyclin H). TFIIH can be resolved into two subcomplexes: core-TFIIH and the cyclin-kinase complex. In addition to its role in transcription, core-TFIIH is also an essential component of the nucleotide excision repair (NER) machinery. This discovery provided a molecular connection between transcription and repair that had been suspected prior to the identification of TFIIH. Holo-TFIIH is essential for transcription, affecting steps before, during, and immediately after initiation.

Recent DNA crosslinking studies have led to two distinctly different models for how XPB catalyzes promoter melting. In one study, XPB was reported to contact DNA both upstream and downstream of the start site, resulting in tight wrapping of the template around the PIC. ATP hydrolysis then induces an XPB conformational change that physically separates the two DNA strands to yield an open promoter complex (Douziech et al., 2000). In the other study, XPB makes more limited DNA contacts, solely downstream of the start site (Kim et al., 2000). These data led to the proposal that XPB catalyzes ATP-dependent rotation of DNA downstream of a rotationally fixed upstream site, thereby functioning as a molecular "wrench" to unwind DNA at the start site. Despite the fundamental differences between these two models for XPB-mediated promoter opening, neither proposes that XPB functions as a classical helicase to unwind template DNA.

Three-dimensional structures of human and yeast TFIIH particles have been deciphered (Figure 1B). Human TFIIH was resolved to 38 Å by single-particle electron microscopy and image processing, revealing a ring-like structure with a central hole whose dimensions are sufficient to accommodate a double-stranded DNA molecule (Schultz et al., 2000). An 18 Å resolution structure was also determined for a TFIIH subcomplex from yeast (Chang and Kornberg, 2000). It is noteworthy that the closed ring visible in human TFIIH is missing from the yeast structure, a discrepancy that could be accounted for by the absence of the yeast counterpart of XPB from the TFIIH subcomplex. The quaternary structure of TFIIH is comparable to other ring-like nucleic acid binding complexes and raises interesting possibilities for how TFIIH binds and remains associated with template DNA.

The Engine: RNAP II and Its CTD

All known cellular RNAPs, including eukaryotic RNAP I, II, and III as well as bacterial RNAPs, are strikingly similar in their subunit composition, amino acid sequence, and function. The largest RNAP II subunit has a carboxy-terminal domain (CTD) consisting of a heptapeptide repeat with the consensus sequence Tyr-Ser-Pro-Thr-Ser-Pro-Ser. The CTD is specific to RNAP II and plays key roles in regulation of transcription initiation and coordination of cotranscriptional mRNA processing events. Four different CTD kinases have been identified, all of which are cyclin regulated (Murray et al., 2001, and references therein). The CTD kinases have specialized functions, exerting either positive or negative effects on transcription by targeting different CTD residues at different stages of the transcription cycle. A CTD phosphatase serves to recycle the phosphorylated form of RNAP II. RNA processing events affected by CTD phosphorylation include 5′ cap addition, splicing, and 3-poly(A) tail addition (see Proudfoot et al., 2002 [this issue of *Cell*]).

Emergence of RNAP II Structure

The size, complexity, and low abundance of RNAP II presented seemingly insurmountable obstacles to determining its three-dimensional structure. The first structure was derived from electron microscopy of 2D crystals, revealing the basic features and contours of the enzyme at 16 Å resolution. A number of hurdles needed to be overcome to obtain crystals for X-ray analysis, which led to a series of major refinements and embellishments. Two of the latest in a series of seminal papers from Kornberg and colleagues describe RNAP II at 2.8 Å resolution in one case and at 3.3 Å resolution in the case of an RNAP II elongation complex (Cramer et al., 2001; Gnatt et al., 2001). Here we highlight some of the mechanistic insights derived from these structures, focusing on RNAP II alone; an analysis of the elongating form of RNAP II was summarized previously (Klug, 2001).

Anatomy of RNAP II

The lion's share of structural studies on yeast RNAP II enlists a 10 subunit version of the 12 subunit enzyme, since two of the subunits (Rpb4 and Rpb7) are present at substoichiometric levels, preventing the isolation of homogenous crystals necessary for high-resolution atomic structures. Unless stated otherwise, structures discussed herein refer to the 10 subunit form of yeast RNAP II.

The overall shape of RNAP II mirrors that of the 3.3 Å resolution bacterial *Thermus aquaticus* RNAP: the bacterial shape is defined as a "crab claw," comparable to the "jaws" of RNAP II (Figures 1C and 1D). The structural similarities between the eukaryotic and bacterial RNAPs are greater than predicted by amino acid sequence similarities (Cramer et al., 2001; Ebright, 2000). Since human and yeast RNAP II share a much higher level of sequence identity that is evenly dispersed among subunits—at both surface and core positions—the yeast structure is certain to be a structural paradigm for all forms of RNAP II. Complementation studies have already demonstrated that most of the human RNAP II subunits can function in place of their yeast counterparts. Homology modeling also shows conservation of the charge distribution between human and yeast RNAP II. The distribution of charge on RNAP II is not only evolutionarily conserved but is amazingly distinct—the surface charge of RNAP II is almost exclusively negative, and the regions that contact DNA are positive.

Comparison of yeast RNAP II and bacterial RNAP reveals that the majority of conserved residues map to the core of the enzyme (including the active center). However, there is no significant conservation of surface residues. The similarities in the core structure suggest analogous catalytic mechanisms, while the differences in surface residues are consistent with the need for the eukaryotic enzyme to make multiple contacts with Mediator, GTFs, or other regulatory proteins not present in bacteria.

The opening between the two jaws is the most easily distinguishable feature of RNAPs, forming a 25 Å diameter cleft or channel where the enzyme clamps onto the template DNA. The presence of a channel with a similar diameter is a unifying feature of RNAPs whose structures have been determined: yeast RNAP II, bacterial RNAP, bacteriophage T7 RNAP, human immunodeficiency virus HIV-1 reverse transcriptase, and even DNA polymerases (Asturias et al., 1997, and references therein). Most of the molecular mass of the two jaws is derived from the two largest subunits: Rpb1 forms most of the lower jaw, and Rpb2 forms most of the upper jaw (Figure 1D).

Some of the finer details, as well as conformational differences, transpired upon comparison of two crystal forms. The two forms, one at 2.8 Å and the other at 3.1 Å resolution, resulted from the use of different crystallization protocols (Cramer et al., 2001). One new feature that came into view in the 2.8 Å form—the presence of two Mg^{2+} ions —helped resolve a conundrum regarding the active site. Lower-resolution forms of RNAP II appeared to have only one Mg^{2+} ion at the active site, inconsistent with the two-metal ion catalyzed mechanism for nucleotide addition proposed for all types of polymerases (Steitz, 1998). With RNAP II now mechanistically aligned with the rest of the polymerases, bacterial RNAP may soon follow suit. To date, only one Mg^{2+} was noted in the 3.3 Å *Thermus aquaticus* RNAP active site. Given the prominent theme of mechanistic conservation in the enzyme core, the presence of two Mg^{2+} ions in bacterial RNAP active centers seems likely once a higher resolution structure becomes available.

Another feature viewed in more detail was the flexible module that pivots over the active center, referred to as the clamp. The clamp constitutes part of one jaw. The majority of the clamp is derived from Rpb1, while the remainder derives from a small portion of Rpb2 (which contributes to a small portion of this jaw) and the Rpb6 amino-terminal tail. The clamp has a fixed range of motion that appears to enable both open and closed positions that alter the size of the cleft. The open position is thought to allow entry of the promoter DNA. When closed, the clamp is proposed to sense the conformation of the DNA-RNA hybrid and separate the DNA and RNA strands upstream of the transcription bubble. Since Rpb6 is phosphorylated, it may influence the position of the clamp (Kayukawa et al., 1999; Kolodziej et al., 1990). The position of Rpb6 within the clamp is consistent with a regulatory role for clamp movement: it is connected to the base of the clamp through a set of five "switches" that serve as control panels for clamp movement.

While the clamp was visible on the 6 Å structure (Fu et al., 1999), three additional mobile modules are now discernible on RNAP II (referred to as the "jaw-lobe," "shelf," and "core" modules). Although the function of these newly identified modules is less clear, it is satisfying to see that there are indeed conduits on the enzyme that can be tweaked. The presence of multiple flexible modules is in agreement with one intrinsic function of RNAP II as the last player in the relay—its ability to sense and respond to signals.

Tracking the CTD

Nascent RNA exits RNAP II through a groove ("groove 1") that originates near the active center and ends at the last ordered residue of Rpb1 in the structure (about 90 amino acids before the CTD begins). These ~90 residues are referred to as the linker to the CTD (Cramer et

al., 2001). The X-ray structure of the CTD and linker region is disordered, precluding its precise visualization in the context of the enzyme. The space occupied by the linker and CTD was approximated by measuring the free space left by four packed crystals. Curiously, the calculated space available for one linker and the CTD could only accommodate a compact, not stretched-out, form of this region comprising the CTD. This observation contrasts with earlier in vitro data demonstrating that phosphorylation caused a major conformational change resulting in a more extended CTD structure (Zhang and Corden, 1991). There may be even more flexible modules within the CTD that are only accessible under certain conditions or stages of transcription. Finally, the fact that the groove exits the enzyme near the CTD is consistent with the coupling of CTD function and RNA processing (see Proudfoot et al., 2002 [this issue of *Cell*]).

RNAP II Subunit Architecture

Although genetic and biochemical experiments defined or suggested certain subunit interactions, it was unclear how the subunits, particularly the small ones, fit together. Upon improving the diffraction of the RNAP II crystals from 6 Å to 3 Å, a polypeptide backbone could be followed for most of the subunits (Cramer et al., 2000). From there, high-resolution structures of two yeast subunits were added, a fit aided by alignment of zinc residues. Earlier functional data pinpointed subunits important for assembly or start site selection; however, the existing data made even more sense in the context of the subunit architecture. Subunits that influenced the position of the initiation site were shown to reside near the DNA channel; those important for assembly were in positions that were consistent with that role; and subunits shown to influence activation were surface exposed and available for interaction with signaling proteins.

The subunit architecture also helped to solidify some predicted functional relationships between subunits, as well as illuminate new relationships not previously apparent. While the functional correlations between Rpb1 and β′ and between Rpb2 and β were readily apparent, other associations took longer to unearth because of very limited sequence similarity. For example, the RNAP II counterpart of the bacterial α^{I} and α^{II} dimer was not clear, although several studies suggested that this role is filled by the Rpb3 and Rpb11 subunits (Tan et al., 2000, and references therein). This was confirmed by the RNAP II backbone structure, showing that Rpb3 and Rpb11 occupy positions comparable to α^{I} and α^{II} in the bacterial enzyme. Even the enigmatic ω subunit of bacterial RNAP is now known to have a counterpart in the RNAP II Rpb6 subunit. This surprising evolutionary link was substantiated through a combination of biochemical, genetic, and structural information (Minakhin et al., 2001). Indeed, Rpb6 and ω occupy comparable positions within their respective structures, wrapping around portions of the enzymes' largest subunit to promote subunit assembly and stability, in agreement with work indicating of role for Rpb6 in assembly (Nouraini et al., 1996). Although Rpb6 is essential for RNAP II activity and cell viability, ω is not required for bacterial viability or for transcription either in vivo or in vitro. The limited sequence similarity between ω and Rpb6 is consistent with these marginal functional parallels.

Missing Pieces: Rpb4-Rpb7

The general location of the Rpb4 and Rpb7 subunits was determined by comparison of the 12 subunit RNAP II structure (at a nominal resolution of 24 Å) to isomorphous crystals from the 10 subunit structure using difference analysis (Asturias et al., 1997; Jensen et al., 1998). The Rpb4-Rpb7 complex occupies a crevice on the surface between Rpb5 in the lower jaw (predominantly Rpb1) and the clamp. Although the overall structures of the two enzymes were highly similar, the 10 subunit RNAP II was in a more open conformation than the wild-type enzyme. This structural data led to the proposition that Rpb4 and Rpb7 may sense entry of DNA into the cleft and induce its closure.

More recent information enabled refinement of the putative role of Rpb4 and Rpb7, and all evidence points toward a role in binding RNA, not DNA. The structure of an Rpb4-Rpb7 complex from *Archaea* revealed the presence of two canonical S1 RNA binding motifs in Rpb7 (Todone et al., 2001). The Rpb4 structure lacks recognizable functional domains but interacts with Rpb7 at the interface between the two S1 domains, suggesting a stabilizing role. These findings are consistent with the earlier discovery (using homology modeling) in Rpb7 of an OB-fold domain of the ribosomal S1 protein subfamily known to bind only ssRNA (Orlicky et al., 2001). The modeled structure accurately predicted function, since the Rpb4-Rpb7 complex could bind ssRNA (as well as ssDNA) in vitro (Orlicky et al., 2001).

More recent information also confirmed the importance of the Rpb4-Rpb7 complex in the stability of the transcription complex. Whole genome expression analysis and biochemical experiments in cells with the 10 subunit RNAP II demonstrated that RNAP II is inactivated at high temperatures and partially active at normal growth temperatures (Maillet et al., 1999; Miyao et al., 2001). In summary, Rpb4 and Rpb7 bind RNA and somehow stabilize the transcription complex. The details remain unclear, but the position of Rpb4-Rpb7 is near groove 1 (the channel for nascent RNA exit), consistent with a role in RNA binding. The proximity of Rpb4-Rpb7 to the flexible clamp and its influence on clamp position in low-resolution structures (open without Rpb4-Rpb7; closed with Rpb4-Rpb7) suggests that this subunit pair may modulate the position of this flexible module (Cramer et al., 2000).

Mediator

Whereas RNAP II and the GTFs are sufficient for promoter-specific transcription in a reconstituted system, additional factors are required for the response to activators. This deficiency led to the discovery of Mediator, a modular complex that transduces both positive and negative regulatory information from gene-specific activators and repressors to the core transcriptional machinery (Hampsey and Reinberg, 1999; Myers and Kornberg, 2000). Unlike RNAP II and the GTFs, Mediator itself is unable to bind specific DNA sequences. Mediator physically interacts with RNAP II, but RNAP II is not a component of Mediator. Mediator also stimulates basal

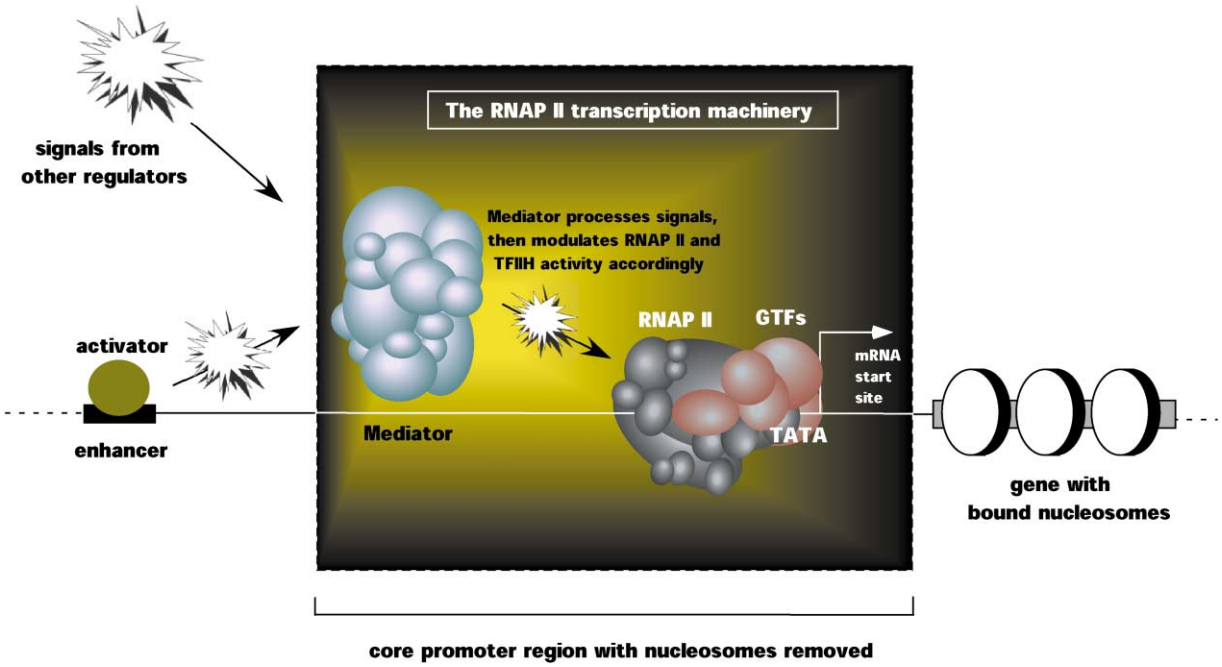


Figure 2. The RNAP II Transcription Machinery

All sizes are approximate. Subunit and transcription factor placement in the respective complexes is not intended to reflect known interactions.

transcription and regulates the TFIIH CTD kinase activity, suggesting that Mediator can function through the CTD. Therefore, our current view of the minimal components of the RNAP II transcription machinery required for regulated transcription initiation includes RNAP II, the GTFs, and Mediator (Figure 2).

Yeast Mediator comprises at least 20 subunits, including Srb proteins, Med proteins, and several other polypeptides originally detected in a variety of genetic screens for transcription factors (Figure 3; Myers and Kornberg, 2000). The Srb proteins were discovered in a genetic screen for suppressors of growth defects associated with RNAP II CTD truncations (Myer and Young, 1998). These include Srb2, Srb4, Srb5, and Srb6, all of which copurified in a complex with RNAP II, referred to as the "RNAP II holoenzyme." The term "Mediator" was first proposed based on its requirement for transcriptional activation in a reconstituted system. The majority of proteins isolated in either the Mediator complex or the RNAP II holoenzyme complex overlapped, providing strong evidence that Mediator functioned through RNAP II. Moreover, a conditional mutation in a component of both complexes (Srb4) nearly eliminated mRNA synthesis at the restrictive temperature, suggesting that Mediator plays a general role in RNAP II transcription.

Following the initial characterization of these complexes in yeast, numerous examples of other RNAP II holoenzyme complexes in yeast and humans were reported. Each new form was purified from the point of view of a tool developed for, or interaction based on, one of the conserved components of the complex. However, most RNAP II holoenzyme complexes do not include an entire complement of the GTFs. Therefore, the term "holoenzyme" does not imply that the complex is sufficient for transcription initiation, as is the case for the classically defined bacterial RNAP holoenzyme. More recently, it has also been demonstrated that while Mediator is sometimes isolated in association with RNAP II (as with the originally isolated RNAP II holoenzyme complex), interaction with RNAP II is not required for Mediator to act as a coactivator in vitro or in vivo (Bhoite et al., 2001; Park et al., 2001). Also, quantification of transcription machinery components argues against a preassembled RNAP II holoenzyme complex (Borggrefe et al., 2001).

The names of individual yeast Mediator proteins are derived from the approaches used for their identification (Figure 3; Myers and Kornberg, 2000). Novel proteins identified in the genetic screen for truncated CTD suppressors were called Srb proteins, and those first discovered following the purification of Mediator were called Med proteins. Most of the remaining Mediator proteins—Gal11, Sin4, Rgr1, Pgd1/Hrs1, Rox3, Nut1, and Nut2—have no unifying nomenclature, since they were discovered in earlier genetic screens for transcription factors (Carlson, 1997). The discovery of these polypeptides in the Mediator not only united these proteins identified in disparate genetic screens into a common biochemical entity, but provided physiological evidence that Mediator functioned as a cofactor for gene regulation in vivo.

Mediator Modules

Various biochemical experiments suggested the existence of distinct subcomplexes of yeast Mediator. One of the more comprehensive approaches used urea to differentially dissociate a recombinant Mediator and to clarify the content of the distinct subcomplexes (Kang et al., 2001). These experiments identified two functionally distinct subcomplexes, the Srb4 module and the Rgr1

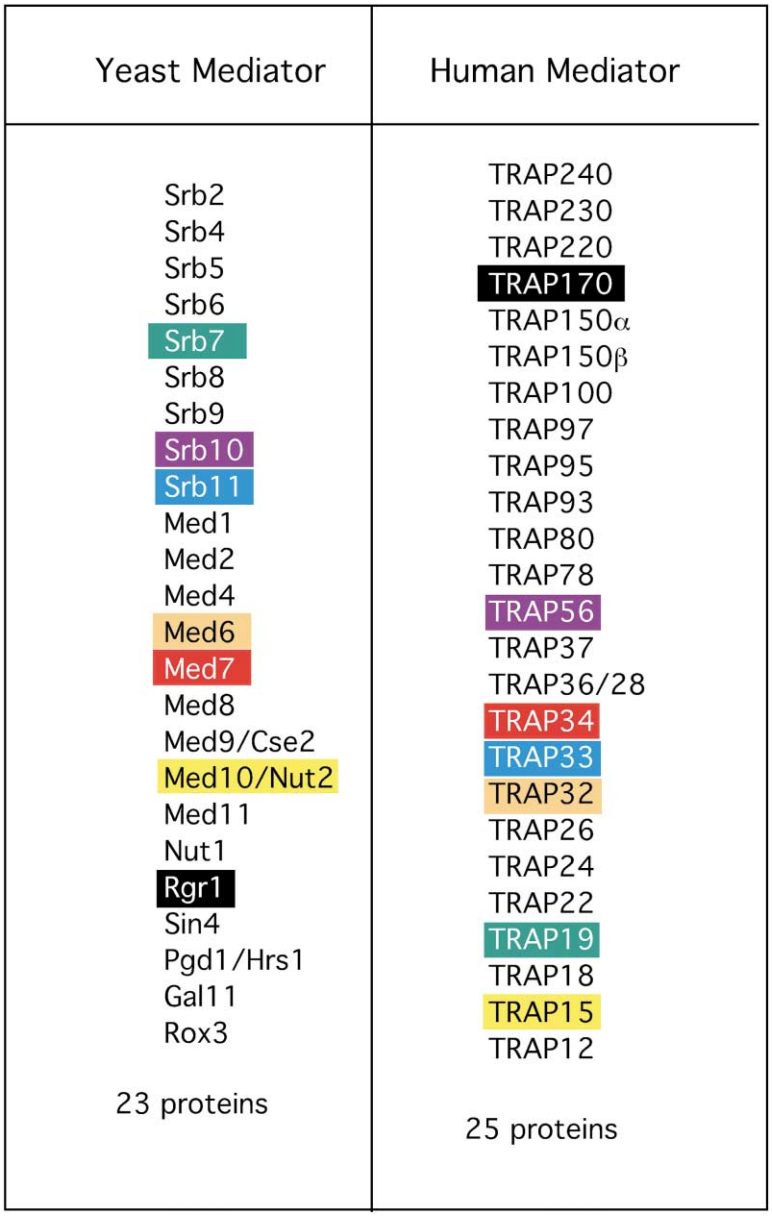


Figure 3. Yeast and Human Mediator Complexes

Subunits of yeast Mediator are grouped and listed by name, not by apparent molecular weight. The example used for human Mediator is the TRAP/SMCC complex; subunit names reflect apparent molecular weight. Evolutionary conservation exists between seven subunits in the two complexes; each related pair is highlighted with matching colors. Subunit similarities adapted from Malik and Roeder (2000).

module. The Srb4 module contains predominantly Srb subunits. Consistent with the genetic identification of *SRB* genes, this module interacts with the RNAP II CTD, along with TBP and TFIIB. The Rgr1 module comprises distinct Gal11 and Med9/10 submodules.

While Mediator appears to be one of the more functionally divergent players in the transcription process, it does exhibit significant evolutionary conservation (Malik and Roeder, 2000; Rachez and Freedman, 2001). However, the conservation is limited to a subset of Mediator subunits. Orthologs for seven of the yeast Mediator components are found in human Mediator preparations (Figure 3). In summary, Mediator is a modular complex that serves as the interface between gene-specific regulatory proteins and the general RNAP II machinery. With the cast of characters in place, the mechanisms by which Mediator serves this function can now be elucidated.

Perplexing Complexes

The composition of human Mediator varies according to the strategy and biochemical tools used for its isolation. However, all approaches were contingent on the general ability of Mediator to either activate or repress transcription. For example, a human Mediator complex was isolated as an enhancer of thyroid hormone receptor function while a related but distinct Mediator complex was isolated as a coactivator of the human transcription factor Sp1. Other variables, such as cell type and growth conditions, may also influence the composition of the purified Mediator complex. Mediator complexes isolated using independent approaches are sometimes identical, but typically have only a subset of common subunits. The modular and variable composition of yeast and human Mediator complexes led to the proposal that Mediator serves as a control panel for the integration of many and diverse regulatory signals that are transmit-

ted to the RNAP II core machinery (Hampsey and Reinberg, 1999).

The jury is still out on the functional significance of the array of Mediators documented. It has been proposed that a minimal module serves as a core Mediator, enabling the cell to custom-design Mediator complexes (Rachez and Freedman, 2001). Each type of Mediator is then formed in response to the signals from transcription factors which, in turn, are summoned in response to the constant fluctuation of the cellular milieu in mature and developing cells. This is the more attractive option because it would elegantly exemplify the exquisite level of responsiveness and efficiency often touted in eukaryotes (especially metazoans). This theory is supported by data in *Caenorhabditis elegans* demonstrating that metazoan-specific Mediator (Med230) is required for expression of stage-specific developmental genes but not two ubiquitously expressed genes. The more mundane, but viable, explanation is that the multiplicity of Mediators simply reflects the biochemical idiosyncrasies of each individual purification approach.

Mediator Architecture

Low-resolution, three-dimensional images of Mediator complexes from yeast, human, and mouse cells have been published (Asturias et al., 1999; Dotson et al., 2000). Mediator forms an elliptical structure that appears to undergo a conformational change in the presence of either RNAP II or the CTD to form a crescent-shaped structure consisting of head, middle, and tail domains (Figure 1E). In a *sin4* mutant the tail domain is missing, but the head and middle domains appear unperturbed. Based on this information and physical interactions among Mediator subunits, the Sin4, Rgr1, and Srb4 submodules have been assigned to the tail, middle, and head domains, respectively (Dotson et al., 2000). The head domain makes the most extensive contacts with RNAP II and is the most structurally conserved domain among the yeast, mouse, and human Mediator complexes. Interestingly, however, mammalian orthologs of Srb2, Srb4, Srb5, and Srb6 polypeptides have not been identified, suggesting that the quaternary structure of head domain is more highly conserved than the primary structures of its individual subunits. The tail domain is the least-conserved structure among the three forms of Mediator, and the *sin4* mutant fails to respond to specific activators while remaining competent for basal transcription. On the other hand, the genes encoding the Srb4 submodule are essential for cell viability and are critical for both basal and activated transcription in vitro, consistent with the apparent direct interaction between the Mediator head domain and RNAP II.

Although the current images of Mediator only achieve a 30–35 Å resolution, it is informative to compare these structures with the biochemical and genetic information. The tail domain is the least highly conserved structure among the three forms of Mediator, and a "tail-less" mutant lacking all of the nonessential components of the Sin4 module fails to respond to specific activators while remaining competent for basal transcription. On the other hand, the genes encoding the Srb4 submodule are essential for cell viability and are critical for both basal and activated transcription in vitro, consistent with the apparent direct interaction between the Mediator head domain and RNAP II.

Concluding Remarks

RNAP II at atomic resolution represents the current high-water mark in the structural analysis of transcription. Solution and crystal structures of the GTFs and Mediator complex at varying degrees of resolution further define this extraordinarily complex machinery, providing insight into the mechanisms that underlie gene expression. These structures are now being exploited not only to interpret and refine the wealth of previous biochemical and genetic information, but to direct future experiments designed to unravel the intricacies of RNAP II transcription.

We have generalized the core RNAP II machinery to include the enzyme itself, the requisite GTFs, and Mediator, the most global of the RNAP II cofactors. While this may be an oversimplification, it serves to organize the highly complex RNAP II machinery into units of discrete function. Nonetheless, there are important caveats to recognize. Foremost, Mediator is unlikely to exist as a single complex, but rather as a set of complexes formed by mixing and matching modular units in response to specific regulators. Also, Mediator is not the only cofactor that transmits regulatory information to RNAP II. TFIID and other TAF complexes can also serve this function, either with, or even in place of, Mediator (Näär et al., 2001). Other gene-specific "mediator" complexes are also likely to be identified. For example, the yeast Paf1 complex transmits regulatory information from the protein kinase C signaling machinery to RNAP II, yet is distinct from both Mediator and TFIID (Chang et al., 1999).

The most conspicuous deficiency in our current view of the RNAP II transcription apparatus is the mechanism of Mediator action. How does Mediator transmit signals between activators or repressors and RNAP II? The current low-resolution structure of Mediator is consistent with the idea that it physically bridges the interaction between regulatory proteins and RNAP II. However, Mediator does not simply facilitate activator-mediated recruitment of the core transcriptional complex to promoter DNA. Indeed, chromatin immunoprecipitation experiments that map the spatial and temporal binding of factors to promoter DNA argue against a role for Mediator in recruitment (Bhoite et al., 2001). Now that the entire cast of characters is in place, a flurry of activity is sure to be directed toward defining the functions of specific Mediator subunits, all with an eye on elucidating the Mediator mechanism. For now, we are simply inspired and humbled by our glimpse into the heart of the RNAP II apparatus.

Acknowledgments

We thank Bo-Shiun Chen and Danny Reinberg for valuable comments on the manuscript. We are also grateful to Francisco Asturias, Seth Darst, Roger Kornberg, and Eva Nogales for generously providing the structural images that appear in Figure 1. N.A.W. is supported by National Institutes of Health grant GM55736 and American Cancer Society grant RPG-97-062-04-GMC. M.H. is supported by NIH grant GM39484.

References

Andel, F., 3rd, Ladurner, A.G., Inouye, C., Tjian, R., and Nogales, E. (1999). Three-dimensional structure of the human TFIID-IIA-IIB complex. Science *286*, 2153–2156.

Asturias, F.J., Meredith, G.D., Poglitsch, C.L., and Kornberg, R.D. (1997). Two conformations of RNA polymerase II revealed by electron crystallography. J. Mol. Biol. *272*, 536–540.

Asturias, F.J., Jiang, Y.W., Myers, L.C., Gustafsson, C.M., and Kornberg, R.D. (1999). Conserved structures of mediator and RNA polymerase II holoenzyme. Science *283*, 985–987.

Bell, S.D., Kosa, P.L., Sigler, P.B., and Jackson, S.P. (1999). Orientation of the transcription preinitiation complex in archaea. Proc. Natl. Acad. Sci. USA *96*, 13662–13667.

Bhoite, L.T., Yu, Y., and Stillman, D.J. (2001). The Swi5 activator recruits the Mediator complex to the HO promoter without RNA polymerase II. Genes Dev. *15*, 2457–2469.

Borggrefe, T., Davis, R., Bareket-Samish, A., and Kornberg, R.D. (2001). Quantitation of the RNA polymerase II transcription machinery in yeast. J. Biol. Chem. *276*, 47150–47153.

Brand, M., Leurent, C., Mallouh, V., Tora, L., and Schultz, P. (1999). Three-dimensional structures of the TAFII-containing complexes TFIID and TFTC. Science *286*, 2151–2153.

Burley, S.K., and Roeder, R.G. (1996). Biochemistry and structural biology of transcription factor IID (TFIID). Annu. Rev. Biochem. *65*, 769–799.

Carlson, M. (1997). Genetics of transcriptional regulation in yeast: connections to the RNA polymerase II CTD. Annu. Rev. Cell Dev. Biol. *13*, 1–23.

Chang, W.H., and Kornberg, R.D. (2000). Electron crystal structure of the transcription factor and DNA repair complex, core TFIIH. Cell *102*, 609–613.

Chang, M., French-Cornay, D., Fan, H.Y., Klein, H., Denis, C.L., and Jaehning, J.A. (1999). A complex containing RNA polymerase II, Paf1p, Cdc73p, Hpr1p, and Ccr4p plays a role in protein kinase C signaling. Mol. Cell. Biol. *19*, 1056–1067.

Cho, E.J., and Buratowski, S. (1999). Evidence that transcription factor IIB is required for a post-assembly step in transcription initiation. J. Biol. Chem. *274*, 25807–25813.

Cramer, P., Bushnell, D.A., Fu, J., Gnatt, A.L., Maier-Davis, B., Thompson, N.E., Burgess, R.R., Edwards, A.M., David, P.R., and Kornberg, R.D. (2000). Architecture of RNA polymerase II and implications for the transcription mechanism. Science *288*, 640–649.

Cramer, P., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 Å resolution. Science *292*, 1863–1876.

Dantonel, J.C., Wurtz, J.M., Poch, O., Moras, D., and Tora, L. (1999). The TBP-like factor: an alternative transcription factor in metazoa? Trends Biochem. Sci. *24*, 335–339.

Dotson, M.R., Yuan, C.X., Roeder, R.G., Myers, L.C., Gustafsson, C.M., Jiang, Y.W., Li, Y., Kornberg, R.D., and Asturias, F.J. (2000). Structural organization of yeast and mammalian mediator complexes. Proc. Natl. Acad. Sci. USA *97*, 14307–14310.

Douziech, M., Coin, F., Chipoulet, J.M., Arai, Y., Ohkuma, Y., Egly, J.M., and Coulombe, B. (2000). Mechanism of promoter melting by the xeroderma pigmentosum complementation group B helicase of transcription factor IIH revealed by protein-DNA photo-cross-linking. Mol. Cell. Biol. *20*, 8168–8177.

Dvir, A., Conaway, J.W., and Conaway, R.C. (2001). Mechanism of transcription initiation and promoter escape by RNA polymerase II. Curr. Opin. Genet. Dev. *11*, 209–214.

Ebright, R.H. (2000). RNA polymerase: structural similarities between bacterial RNA polymerase and eukaryotic RNA polymerase II. J. Mol. Biol. *304*, 687–698.

Evans, R., Fairley, J.A., and Roberts, S.G. (2001). Activator-mediated disruption of sequence-specific DNA contacts by the general transcription factor TFIIB. Genes Dev. *15*, 2945–2949.

Faitar, S.L., Brodie, S.A., and Ponticelli, A.S. (2001). Promoter-specific shifts in transcription initiation conferred by yeast TFIIB mutations are determined by the sequence in the immediate vicinity of the start sites. Mol. Cell. Biol. *21*, 4427–4440.

Forget, D., Robert, F., Grondin, G., Burton, Z.F., Greenblatt, J., and Coulombe, B. (1997). RAP74 induces promoter contacts by RNA polymerase II upstream and downstream of a DNA bend centered on the TATA box. Proc. Natl. Acad. Sci. USA *94*, 7150–7155.

Freiman, R.N., Albright, S.R., Zheng, S., Sha, W.C., Hammer, R.E., and Tjian, R. (2001). Requirement of tissue-selective TBP-associated factor TAFII105 in ovarian development. Science *293*, 2084–2087.

Fu, J., Gnatt, A.L., Bushnell, D.A., Jensen, G.J., Thompson, N.E., Burgess, R.R., David, P.R., and Kornberg, R.D. (1999). Yeast RNA polymerase II at 5 Å resolution. Cell *98*, 799–810.

Gaiser, F., Tan, S., and Richmond, T.J. (2000). Novel dimerization fold of RAP30/RAP74 in human TFIIF at 1.7 Å resolution. J. Mol. Biol. *302*, 1119–1127.

Gnatt, A.L., Cramer, P., Fu, J., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. Science *292*, 1876–1882.

Green, M.R. (2000). TBP-associated factors (TAF_{II}s): multiple, selective transcriptional mediators in common complexes. Trends Biochem. Sci. *25*, 59–63.

Groft, C.M., Uljon, S.N., Wang, R., and Werner, M.H. (1998). Structural homology between the Rap30 DNA-binding domain and linker histone H5: implications for preinitiation complex assembly. Proc. Natl. Acad. Sci. USA *95*, 9117–9122.

Grossmann, J.G., Sharff, A.J., O'Hare, P., and Luisi, B. (2001). Molecular shapes of transcription factors TFIIB and VP16 in solution: implications for recognition. Biochemistry *40*, 6267–6274.

Hahn, S. (1998). The role of TAFs in RNA polymerase II transcription. Cell *95*, 579–582.

Hampsey, M., and Reinberg, D. (1999). RNA polymerase II as a control panel for multiple coactivator complexes. Curr. Opin. Genet. Dev. *9*, 132–139.

Hawkes, N.A., and Roberts, S.G. (1999). The role of human TFIIB in transcription start site selection in vitro and in vivo. J. Biol. Chem. *274*, 14337–14343.

Hawkes, N.A., Evans, R., and Roberts, S.G. (2000). The conformation of the transcription factor TFIIB modulates the response to transcriptional activators in vivo. Curr. Biol. *10*, 273–276.

Hayashi, F., Ishima, R., Liu, D., Tong, K.I., Kim, S., Reinberg, D., Bagby, S., and Ikura, M. (1998). Human general transcription factor TFIIB: conformational variability and interaction with VP16 activation domain. Biochemistry *37*, 7941–7951.

Holmes, M.C., and Tjian, R. (2000). Promoter-selective properties of the TBP-related factor TRF1. Science *288*, 867–870.

Jensen, G.J., Meredith, G., Bushnell, D.A., and Kornberg, R.D. (1998). Structure of wild-type yeast RNA polymerase II and location of Rpb4 and Rpb7. EMBO J. *17*, 2353–2358.

Kamada, K., De Angelis, J., Roeder, R.G., and Burley, S.K. (2001). Crystal structure of the C-terminal domain of the RAP74 subunit of human transcription factor IIF. Proc. Natl. Acad. Sci. USA *98*, 3115–3120.

Kang, J.S., Kim, S.H., Hwang, M.S., Han, S.J., Lee, Y.C., and Kim, Y.J. (2001). The structural and functional organization of the yeast mediator complex. J. Biol. Chem. *276*, 42003–42010.

Kayukawa, K., Makino, Y., Yogosawa, S., and Tamura, T. (1999). A serine residue in the N-terminal acidic region of rat RPB6, one of the common subunits of RNA polymerases, is exclusively phosphorylated by casein kinase II in vitro. Gene *234*, 139–147.

Kim, T.K., Lagrange, T., Wang, Y.H., Griffith, J.D., Reinberg, D., and Ebright, R.H. (1997). Trajectory of DNA in the RNA polymerase II transcription preinitiation complex. Proc. Natl. Acad. Sci. USA *94*, 12268–12273.

Kim, T.K., Ebright, R.H., and Reinberg, D. (2000). Mechanism of ATP-dependent promoter melting by transcription factor IIH. Science *288*, 1418–1422.

Klug, A. (2001). Structural biology. A marvellous machine for making messages. Science *292*, 1844–1846.

Kolodziej, P.A., Woychik, N., Liao, S.M., and Young, R.A. (1990). RNA polymerase II subunit composition, stoichiometry, and phosphorylation. Mol. Cell. Biol. *10*, 1915–1920.

Lagrange, T., Kapanidis, A.N., Tang, H., Reinberg, D., and Ebright, R.H. (1998). New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. Genes Dev. *12*, 34–44.

Leuther, K.K., Bushnell, D.A., and Kornberg, R.D. (1996). Two-dimensional crystallography of TFIIB- and IIE-RNA polymerase II complexes: implications for start site selection and initiation complex formation. Cell *85*, 773–779.

Maillet, I., Buhler, J.M., Sentenac, A., and Labarre, J. (1999). Rpb4p is necessary for RNA polymerase II activity at high temperature. J. Biol. Chem. *274*, 22586–22590.

Malik, S., and Roeder, R.G. (2000). Transcriptional regulation through Mediator-like coactivators in yeast and metazoan cells. Trends Biochem. Sci. *25*, 277–283.

Minakhin, L., Bhagat, S., Brunning, A., Campbell, E.A., Darst, S.A., Ebright, R.H., and Severinov, K. (2001). Bacterial RNA polymerase subunit ω and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. Proc. Natl. Acad. Sci. USA *98*, 892–897.

Miyao, T., Barnett, J.D., and Woychik, N.A. (2001). Deletion of the RNA polymerase subunit RPB4 acts as a global, not stress-specific, shut-off switch for RNA polymerase II transcription at high temperatures. J. Biol. Chem. *276*, 46408–46413.

Murray, S., Udupa, R., Yao, S., Hartzog, G., and Prelich, G. (2001). Phosphorylation of the RNA polymerase II carboxy-terminal domain by the Bur1 cyclin-dependent kinase. Mol. Cell. Biol. *21*, 4089–4096.

Myer, V.E., and Young, R.A. (1998). RNA polymerase II holoenzymes and subcomplexes. J. Biol. Chem. *273*, 27757–27760.

Myers, L.C., and Kornberg, R.D. (2000). Mediator of transcriptional regulation. Annu. Rev. Biochem. *69*, 729–749.

Näär, A.M., Lemon, B.D., and Tjian, R. (2001). Transcriptional coactivator complexes. Annu. Rev. Biochem. *70*, 475–501.

Nikolov, D.B., and Burley, S.K. (1997). RNA polymerase II transcription initiation: a structural view. Proc. Natl. Acad. Sci. USA *94*, 15–22.

Nouraini, S., Archambault, J., and Friesen, J.D. (1996). Rpo26p, a subunit common to yeast RNA polymerases, is essential for the assembly of RNA polymerases I and II and for the stability of the largest subunits of these enzymes. Mol. Cell. Biol. *16*, 5985–5996.

Okuda, M., Watanabe, Y., Okamura, H., Hanaoka, F., Ohkuma, Y., and Nishimura, Y. (2000). Structure of the central core domain of TFIIEβ with a novel double-stranded DNA-binding surface. EMBO J. *19*, 1346–1356.

Orlicky, S.M., Tran, P.T., Sayre, M.H., and Edwards, A.M. (2001). Dissociable Rpb4-Rpb7 subassembly of RNA polymerase II binds to single-strand nucleic acid and mediates a post-recruitment step in transcription initiation. J. Biol. Chem. *276*, 10097–10102.

Orphanides, G., LaGrange, T., and Reinberg, D. (1996). The general initiation factors of RNA polymerase II. Genes Dev. *10*, 2657–2683.

Park, J.M., Werner, J., Kim, J.M., Lis, J.T., and Kim, Y.J. (2001). Mediator, not holoenzyme, is directly recruited to the heat shock promoter by HSF upon heat shock. Mol. Cell *8*, 9–19.

Proudfoot, N., Furger, A., and Dye, M.J. (2002). Integrating mRNA processing with transcription. Cell *108*, this issue, 501–512.

Rachez, C., and Freedman, L.P. (2001). Mediator complexes and transcription. Curr. Opin. Cell Biol. *13*, 274–280.

Ranish, J.A., Yudkovsky, N., and Hahn, S. (1999). Intermediates in formation and activity of the RNA polymerase II preinitiation complex: holoenzyme recruitment and a postrecruitment role for the TATA box and TFIIB. Genes Dev. *13*, 49–63.

Reinberg, D., Orphanides, G., Ebright, R., Akoulitchev, S., Carcamo, J., Cho, H., Cortes, P., Drapkin, R., Flores, O., Ha, I., et al. (1998). The RNA polymerase II general transcription factors: past, present, and future. Cold Spring Harb. Symp. Quant. Biol. *63*, 83–103.

Robert, F., Douziech, M., Forget, D., Egly, J.M., Greenblatt, J., Burton, Z.F., and Coulombe, B. (1998). Wrapping of promoter DNA around the RNA polymerase II initiation complex induced by TFIIF. Mol. Cell *2*, 341–351.

Schultz, P., Fribourg, S., Poterszman, A., Mallouh, V., Moras, D., and Egly, J.M. (2000). Molecular structure of human TFIIH. Cell *102*, 599–607.

Selleck, W., Howley, R., Fang, Q., Podolny, V., Fried, M.G., Buratowski, S., and Tan, S. (2001). A histone fold TAF octamer within the yeast TFIID transcriptional coactivator. Nat. Struct. Biol. *8*, 695–700.

Steitz, T. (1998). A mechanism for all polymerases. Nature *391*, 231–232.

Tan, Q., Linask, K.L., Ebright, R.H., and Woychik, N.A. (2000). Activation mutants in yeast RNA polymerase II subunit RPB3 provide evidence for a structurally conserved surface required for activation in eukaryotes and bacteria. Genes Dev. *14*, 339–348.

Todone, F., Brick, P., Werner, F., Weinzierl, R.O., and Onesti, S. (2001). Structure of an archaeal homolog of the eukaryotic RNA polymerase II RPB4/RPB7 complex. Mol. Cell *8*, 1137–1143.

Tsai, F.T., and Sigler, P.B. (2000). Structural basis of preinitiation complex assembly on human Pol II promoters. EMBO J. *19*, 25–36.

Verrijzer, C.P., and Tjian, R. (1996). TAFs mediate transcriptional activation and promoter selectivity. Trends Biochem. Sci. *21*, 338–342.

Weil, P.A., Luse, D.S., Segall, J., and Roeder, R.G. (1979). Selective and accurate initiation of transcription at the Ad2 major late promotor in a soluble system dependent on purified RNA polymerase II and DNA. Cell *18*, 469–484.

Woychik, N.A. (1998). Fractions to functions: RNA polymerase II thirty years later. Cold Spring Harb. Symp. Quant. Biol. *63*, 311–317.

Zhang, J., and Corden, J.L. (1991). Phosphorylation causes a conformational change in the carboxy-terminal domain of the mouse RNA polymerase II largest subunit. J. Biol. Chem. *266*, 2297–2302.

Zhang, D., Penttila, T.L., Morris, P.L., Teichmann, M., and Roeder, R.G. (2001). Spermiogenesis deficiency in mice lacking the Trf2 gene. Science *292*, 1153–1155.

Zhu, W.L., Zeng, Q.D., Colangelo, C.M., Lewis, L.M., Summers, M.F., and Scott, R.A. (1996). The N-terminal domain of TFIIB from *Pyrococcus furiosus* forms a zinc ribbon. Nat. Struct. Biol. *3*, 122–124.